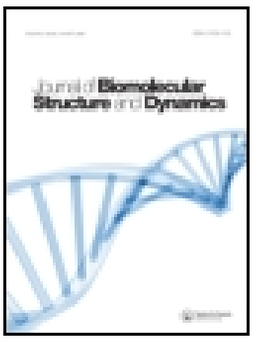

Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

The mechanism of high affinity pentasaccharide binding to antithrombin, insights from Gaussian accelerated molecular dynamics simulations

Gábor Balogh, István Komáromi & Zsuzsanna Bereczky

To cite this article: Gábor Balogh, István Komáromi & Zsuzsanna Bereczky (2019): The mechanism of high affinity pentasaccharide binding to antithrombin, insights from Gaussian accelerated molecular dynamics simulations, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2019.1688194

To link to this article: https://doi.org/10.1080/07391102.2019.1688194

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2019.1688194

Taylor & Francis
Taylor & Francis Group

The mechanism of high affinity pentasaccharide binding to antithrombin, insights from Gaussian accelerated molecular dynamics simulations

Gábor Balogh[a,b], István Komáromi[†a] and Zsuzsanna Bereczky[a]

[a]Division of Clinical Laboratory Science, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary; [b]Kálmán Laki Doctoral School, University of Debrecen, Debrecen, Hungary

Communicated by Ramaswamy H. Sarma

ABSTRACT

The activity of antithrombin (AT), a serpin protease inhibitor, is enhanced by heparin and heparin analogs against its target proteases, mainly thrombin, factors Xa and IXa. Considerable amount of information is available on the multistep mechanism of the heparin pentasaccharide binding and conformational activation. However, much of the details were inferred from 'static' structures obtained by X-ray diffraction. Moreover, limited information is available for the early steps of binding mechanism other than kinetic studies with various ligands. To gain insights into these processes, we performed enhanced sampling molecular dynamics (MD) simulations using the Gaussian Accelerated Molecular Dynamics (GAMD) method, applied previously in drug binding studies. We were able to observe the binding of the pentasaccharide idraparinux to a 'non-activated' AT conformation in two separate trajectories with low root mean square deviation (RMSD) values compared to X-ray structures of the bound state. These trajectories along with further simulations of the AT-pentasaccharide complex provided insights into the mechanisms of multiple conformational transitions, including the expulsion of the hinge region, the extension of helix D and the conformational behavior of the reactive center loop (RCL). We could also confirm the high stability of helix P in non-activated AT conformations, such states might play an important role in heparin binding. 'Generalized correlation' matrices revealed possible paths of allosteric signal propagation to the binding sites for the target proteases, factors Xa and IXa. Enhanced MD simulations of ligand binding to AT may assist the design of new anticoagulant drugs.

ARTICLE HISTORY
Received 10 September 2019
Accepted 29 October 2019

KEYWORDS
Antithrombin; heparin pentasaccharide; ligand binding; binding mechanism; conformational change; molecular dynamics; Gaussian accelerated molecular dynamics

1. Introduction

Antithrombin (AT), a plasma glycoprotein belonging to the serpin superfamily, is one of the key regulators of coagulation cascade (Quinsey, Greedy, Bottomley, Whisstock, & Pike, 2004). Members of the serpin superfamily are found in all branches of life (Rau, Beaulieu, Huntington, & Church, 2007). They include inhibitors of serine and cysteine proteases, as well as non-inhibitory proteins (Huntington, 2011; Whisstock & Bottomley, 2006). The serpins share a common tertiary structure, consisting of eight alpha-helices and three beta-sheets as well as a solvent-exposed reactive center loop (RCL) in their 'native' form (Huntington, 2006; Whisstock & Bottomley, 2006). The protease inhibition mechanism of the inhibitory serpins involves the formation of a covalent complex with the protease (Huntington, 2011; Whisstock & Bottomley, 2006). The 'trapping' of the protease is accompanied by significant conformational changes of the serpin, most importantly the insertion of the N-terminal part of the partially cleaved RCL into beta-sheet A (Huntington, 2006; Whisstock & Bottomley, 2006).

Antithrombin inhibits multiple blood coagulation factors with serine protease activity. Its main targets are thrombin and the activated blood coagulation factors IX and X (FIXa and FXa, respectively) (Muszbek, Bereczky, Kovács, & Komáromi, 2010; Rau et al., 2007). In its circulating, 'native' form, AT inhibits circulating proteases only at relatively low reaction rates. The reaction rates of protease inhibition by antithrombin are enhanced by heparinoids (Olson, Richard, Izaguirre, Schedin-Weiss, & Gettins, 2010; Rau et al., 2007). Two different mechanisms for the greater inhibitory activity have been proposed (Olson et al., 2010). In the first mechanism, binding of the ligand's specific pentasaccharide sequence to AT causes conformational changes. This 'conformational' activation plays a critical role in case of FIXa and FXa, but its significance is limited in case of thrombin (Olson et al., 2010; Rau et al., 2007). Alternatively, longer heparin chains can form 'bridges' between antithrombin and the protease, thus accelerating the formation of complexes. The bridging mechanism is essential for enhanced thrombin inhibitory activity (Olson et al., 2010). Heparin and its analogues are important anticoagulant drugs, and their in vitro effects on the inactivation of thrombin and FXa have been subject of numerous studies. It has been proposed that AT

CONTACT Zsuzsanna Bereczky zsbereczky@med.unideb.hu University of Debrecen, Faculty of Medicine, Department of Laboratory Medicine, Division of Clinical Laboratory Science, 98, Nagyerdei krt., Debrecen, H-4032, Hungary
[†]Passed away on 4 August 2017.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07391102.2019.1688194

could bind heparan sulfate chains at the endothelial wall, therefore limiting blood coagulation when the blood vessel wall is intact (Dinarvand, Yang, Villoutreix, & Rezaie, 2018; Olson et al., 2010).

The conformation changes required for AT activation have been subject of several studies. A three-step model has been proposed for the heparin pentasaccharide binding and activation, involving two consecutive sets of conformational changes (Olson et al., 2010; Roth et al., 2015; Schedin-Weiss, Richard, & Olson, 2010). This mechanism is supported by X-ray diffraction structures (Johnson & Huntington, 2003), kinetic studies of heparin binding and protease inactivation (Schedin-Weiss et al., 2010) as well as by experiments with engineered AT variants (Roth et al., 2015). In the first step, the pentasaccharide forms relatively weak interactions with the heparin binding site of AT, which then triggers conformation changes, most significantly in the regions close to the binding site (Desai, Petitou, Björk, & Olson, 1998; Olson et al., 2010). The second step involves rearrangements in the binding site as well as the hydrophobic core of the protein (Langdown, Belzar, Savory, Baglin, & Huntington, 2009), while the expulsion of the hinge region and the C-terminal extension of D helix in the third step leads to an even higher protease inhibitory activity (Belzar, Zhou, Carrell, Gettins, & Huntington, 2002; Langdown et al., 2009; Olson et al., 2010). X-ray diffraction structures of 'native' and fully activated AT, as well as a structure showing AT in an intermediate activated state have been published (Johnson & Huntington, 2003; McCoy, Pei, Skinner, Abrahams, & Carrell, 2003). In this intermediate structure, most of the conformational changes involving the heparin binding site (with the exception of helix D extension) and the hydrophobic core of the protein have occurred (Johnson & Huntington, 2003). However, all information related to the first step of pentasaccharide binding was obtained by means of kinetic studies, no X-ray structure exists for such an AT conformation.

Although the comparison of multiple AT X-ray diffraction structures representing different conformations added useful pieces of information to understand the allosteric mechanisms, the X-ray structures still show a 'static' image of the various states. Another limitation of these studies is that crystal contacts can affect the conformation of the crystallized protein. The size of AT still precludes the detailed structure-function studies using nuclear magnetic resonance spectroscopy (NMR), especially at an atomic resolution. In contrast computational methods, including molecular dynamics (MD) simulations have proven useful for studying conformational changes and other biologically important processes in proteins and other biomolecules at an atomistic level. Several studies have been published in which MD-based methods were applied to processes involving serpins, including AT. Notable examples for such studies include modelling the transition into latent conformation in plasminogen activator inhibitor 1 at atomistic level (Cazzolli et al., 2014), the elucidation of the pathway of RCL insertion in α1-antitrypsin (Andersen et al., 2017), and protein folding studies aimed at explaining the misfolding of mutated serpins (Wang et al., 2018). Specifically, in case of AT, the role of Asn135 glycosylation (Pol-Fachin, Franco Becker, Almeida Guimarães, & Verli, 2011) as well as the RCL dynamics and the conformation of the critical Arg393 side chain was investigated (Tóth, Fekete, Balogh, Bereczky, & Komáromi, 2015). In a recent study, the binding of D-myo-inositol-3,4,5,6-tetrakisphosphate, a ligand with a scaffold different from heparinoids was studied and signs of allosteric activation such as higher flexibility of RCL were detected (Arantes, Pérez-Sánchez, & Verli, 2018).

However, one of the main limitations of 'conventional' MD remains its high computational cost. This makes the modelling of the so-called 'rare events' challenging. To overcome such 'limitations,' advanced sampling techniques have been developed (Bernardi, Melo, & Schulten, 2015). Examples of such advanced-sampling methods are replica exchange (Sugita & Okamoto, 1999), umbrella sampling (Kästner, 2011), metadynamics (Barducci, Bonomi, & Parrinello, 2011), adaptively biased MD (Babin, Karpusenka, Moradi, Roland, & Sagui, 2009) and simulated annealing (Brünger, Adams, & Rice, 1997). The Gaussian Accelerated Molecular Dynamics (GAMD) method, developed by Miao et al., offers advanced sampling without requiring pre-defined reaction coordinates or collective variables (Miao, Feher, & McCammon, 2015). One of the main advantages of GAMD compared to the related Accelerated Molecular Dynamics (AMD) technique (Hamelberg, Mongan, & McCammon, 2004) is that it allows free-energy calculations using reweighting that is often not possible in case of AMD due to energetic noise (Miao et al., 2014, 2015). The GAMD method has proven useful in modelling of drug-binding pathways to multiple targets (Bhattarai & Miao, 2018), including G-protein coupled receptors (Liao & Wang, 2019; Miao, Bhattarai, Nguyen, Christopoulos, & May, 2018; Miao & McCammon, 2016) and the HIV protease (Miao, Huang, Walker, McCammon, & Chang, 2018). GAMD provides highly enhanced conformational sampling and has the potential to provide realistic ligand binding pathway as it does not depend on pre-defined reaction coordinates (Bhattarai & Miao, 2018).

In the present work, we performed GAMD simulations to investigate the 'early' events of the AT – pentasaccharide binding, with idraparinux (Petitou & van Boeckel, 2004) chosen as a pentasaccharide ligand. Only indirect information exists for these early steps, based on kinetic studies with modified pentasaccharides. The trajectories obtained in our simulations revealed possible pathways concerning the binding of this molecule to AT. Furthermore, our simulations provided insights into the AT conformation changes triggered by pentasaccharide binding.

2. Methods

2.1. System preparation

To study the binding of pentasaccharide to AT, we created three model systems: an AT conformation without allosteric activation, the same AT conformation with an added pentasaccharide and AT-pentasaccharide complex. For the first model system the 1T1F X-ray diffraction structure of native AT was used (Johnson et al., 2006). Here, the structure of

native AT was determined in a monomeric state, without a bound pentasaccharide or a longer heparin chain.

The second model based on the same X-ray structure but contained an added pentasaccharide (idraparinux). In a previous study, we have already investigated the conformation of idraparinux in explicit solvent, using both Gaussian Accelerated Molecular Dynamics and 'conventional' molecular dynamics simulations (Balogh et al., 2019). The initial structure of the idraparinux molecule was built using the 'NTP' ligand found in the AT-pentassacharide complex X-ray diffraction structure 1NQ9. Missing hydrogen atoms were added using the OpenBabel software (O'Boyle et al., 2011). In case of this 'not activated' AT with added ligand system, the initial position of the pentasaccharide was determined using least squares fitting of the AT conformation found in 1T1F on the AT-pentasaccharide complex structure 1NQ9. After the fitting, the ligand was shifted by 12 Å along the vector connecting the centers of masses between AT and the ligand. This resulted in a structure without any contact between AT and pentasaccharide.

An X-ray diffraction structure of the partially activated AT with a bound pentasaccharide (1NQ9) was used as the initial structure for the AT-ligand complex system (third model system) (Johnson & Huntington, 2003). Idrapariux was used as ligand in this structure as well.

The 'engineered' amino acid substitutions in the X-ray structures were mutated back and the missing loops in the protein were built using MODELLER (Sali & Blundell, 1993). The glycosylation of AT was modelled by adding disaccharides consisting of two N-acetylglucosamine units to the AT residues Asn96, Asn155 and Asn192. The three glycosylation sites corresponded to those found in β-antithrombin (Brennan, George, & Jordan, 1987; Quinsey et al., 2004). The reason for using such truncated oligosaccharides was that the appropriate sampling of the large and highly flexible glycans would have been computationally very demanding.

The CHARMM carbohydrate force field was chosen for the ligand (Guvench et al., 2011; Mallajosyula, Guvench, Hatcher, & Mackerell, 2012), because this force field was found most suitable for reproducing NMR data for the same molecule in a previous study (Balogh et al., 2019). For the AT protein, the CHARMM36m force field was applied (Huang et al., 2017). The topology generation and the solvation steps of the system preparation were performed using the CHARMM-GUI webserver (Jo, Kim, Iyer, & Im, 2008; Lee et al., 2016; Park et al., 2017). The model systems were solvated in a cubic box. The ionic strength was set to approximately 0.15 M by adding the appropriate number of Na^+ and Cl^- ions.

2.2. Molecular dynamics and Gaussian accelerated molecular dynamics simulations

We performed multiple 'conventional' MD and GAMD simulations on the three model systems mentioned in the previous section. The AMBER16 pmemd.cuda software was used in all cases (Case et al., 2016; Le Grand, Götz, & Walker, 2013; Salomon-Ferrer, Götz, Poole, Le Grand, & Walker, 2013). The simulations in the present work are summarized in Table S1 (supplementary material).

Before the MD and advanced-sampling MD simulations, energy minimization was performed. In the first 2000 steps the steepest descents algorithm was applied, with position restraint of 5 kcal/mol/Å on all non-solvent atoms. This was followed by another minimization of 2000 steps without restraints (500 steps steepest descents, 1500 steps conjugate gradients). After minimization, the system was heated to 310 K in a 2 ns Langevin dynamics simulation. The heating was followed by a 2 ns pressure equilibration MD using the Berendsen barostat (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984). Position restraints of 2 kcal/mol/Å were used on the protein and ligand atoms in the heating and pressure equilibration steps. A 150-ns molecular dynamics simulation under NPT conditions was performed in all cases before the 'production' MD or GAMD. For the 'non-activated' AT (1T1F) with added pentasaccharide system, four independent 150 ns conventional MD simulations were performed, labelled A–D (Table S1). The 'stochastic' thermostat used in the simulations resulted in divergent trajectories even in the cases where identical initial conformation was used. Two 'starting' conformations, A and B, were chosen for further GAMD and conventional MD simulation, based on the position of the pentasaccharide (see RMSD calculations in the 'MD trajectory analysis' section and Figure S1, supplementary material).

The 'equilibration' phase of the GAMD simulations was 60 ns long (Miao et al., 2015). The protocol of the equilibration step was similar to that used in our previous study of idraparinux (Balogh et al., 2019). The first part of the simulation was 10 ns conventional MD, in which energy data was collected in the last 6 ns. In the following 50 ns part of the simulations, GAMD boost potential was added and the GAMD parameters were continuously updated (except the first 5 ns). A 'dual-boost' GAMD scheme was applied and the upper limits for both the potential and dihedral energy boost SDs (σ_0) were set to 6.0 kcal/mol. Starting from the equilibrated structures, two independent GAMD simulations were performed in all cases: 'native' AT without ligand, 'partially' activated AT with ligand, as well as the two different starting structures of 'native' AT with added idraparinux, obtained by conventional MD. In the latter case, the simulations were labelled A1, A2, B1 and B2 (Table S1). The 'production' simulations were performed under NVT conditions at 310 K using a Langevin thermostat with collision frequency of 2 ps^{-1}. The integration time step was 2 fs. The long-range electrostatic interactions were treated using the PME method (Darden, York, & Pedersen, 1993). The Coulomb cutoff was set to 12 Å. A Lennard-Jones force switch was used between 10 and 12 Å. Bonds between heavy atoms and hydrogens were constrained using the SHAKE algorithm (Ryckaert, Ciccotti, & Berendsen, 1977). The atom coordinates and the GAMD energies were saved every 0.1 ps to make free energy calculations using GAMD reweighting possible.

Two of the four GAMD 'ligand binding' simulations, where, as compared to the 1NQ9 structure, low RMSD values of the ligand were observed, were expanded to 1000 ns.

(Additional details of the RMSD calculations are found in the 'Trajectory analysis' section.) 600 ns equilibrium MD simulations were also performed on all model systems as well. The equilibrium MDs were performed under NPT conditions, but all other parameters were identical to those used in GAMD.

2.3. Trajectory analysis

The CPPTRAJ program was used for trajectory analysis, including calculation of distances, fluctuations and RMSDs (Roe & Cheatham, 2013). With the exception of GAMD reweighting, one snapshot was taken from every 100 ps of the simulations for analysis. The 'pentasaccharide RMSD' was calculated using the following two-step method: alpha carbon atoms in amino acids 6–26, 39–133, 137–355, 361–377, 402–431 of AT were fitted to the 1NQ9 X-ray diffraction structure, and then the RMSD between the ring atoms and the four glycosidic oxygen atoms was determined. Therefore, the conformation of the rotatable side chains of the molecule was not taken into account. Protein secondary structure was analyzed using the DSSP method developed by Kabsch and Sander (1983).

Representative conformations for the observed AT conformation changes were obtained by RMSD-based clustering using the K-means algorithm, implemented in CPPTRAJ (Figure S2, supplementary material). During the cluster analysis, fitting was performed first, using the same atoms as in the pentasaccharide RMSD calculation. The RMSDs for all non-hydrogen atoms in the regions analyzed (Helix D: residues 130-140, Hinge region: 375-385, RCL: 382-401) were used as 'distance metric.' A different approach was used in case of helix P: the trajectory was split into two parts at a point when a conformational change occurred, and the lowest RMSD frame compared to the average structures were used.

GAMD reweighting was performed using the cumulant expansion method as implemented in the PyReweighting program (Miao et al., 2014, 2015). Two-dimensional plots from the reweighting results were created with the DPlot software (https://www.dplot.com/index.htm). Allosteric pathways were studied using the 'generalized correlation' or 'mutual information' method, as implemented by Lange and Grubmüller in the g_correlation program (Lange & Grubmüller, 2005), based on an earlier work by Ichiye and Karplus (1991). The UCSF Chimera software package was used for visualization of protein structures (Pettersen et al., 2004).

3. Results and discussion

3.1. Binding of the pentasaccharide ligand to antithrombin

A three-step model has been proposed for the heparin pentasaccharide binding and the subsequent conformational changes, supported by kinetic studies and an X-ray diffraction structures showing AT in various pentasaccharide-bound and 'native' states (Olson et al., 2010; Schedin-Weiss et al., 2010). However, the mechanism for the earliest steps of binding was mostly inferred from kinetic data (Olson et al., 2010), no X-ray diffraction structure is available that would correspond to a weak AT-pentasaccharide complex. In all available 'complex' structures, AT exhibits either an intermediate or fully activated conformation (Johnson & Huntington, 2003; McCoy et al., 2003).

To study the early events of the binding of the pentasaccharide at an atomistic level, we performed GAMD as well as equilibrium MD simulations. The possible binding modes of the ligands explored by the simulations are depicted in Figure S3 (supplementary material). Here, one conformation is shown for every 50 ns simulation time, colored according to the RMSD compared to the 1NQ9 X-ray diffraction structure. From our four simulations in which 'native' AT with an added pentasaccharide was investigated (labelled A1, A2, B1 and B2), we observed pentasaccharide binding in a conformation comparable to X-ray diffraction in two of the four trajectories, A1 and B1. This was demonstrated by RMSD values < 2.5 Å for a significant period of time. In case of the B1 system, this happened mostly in the 450–800 ns time interval, while in the A1 trajectory, the change occurred at about 700 ns, and the ligand remained largely in the same position until the end of simulation. However, in the other two GAMD trajectories, as well as the equilibrium MD simulations with the added ligand, this ligand-binding conformation was not observed (see Figure S4, supplementary material).

The lowest-RMSD conformation among 1000 extracted frames from these two trajectories is shown in Figure 1a. The lowest RMSD snapshot from the last 100 ns part of the 1 μs long 1NQ9 simulation, representing a similar binding state, is also depicted. In Figure 1b, the three structures were superimposed so that their conformations can be compared to the 1NQ9 structure. As expected, the positions of the pentasaccharide in the three cases are remarkably similar to that found in the X-ray structure. The differences were slightly higher in the GH end of the idraparinux molecule than in the DEF part.

To gain insight into the mechanism of the pentasaccharide binding, we analyzed the RMSD values also for rings D, F and H separately. These correspond to the non-reducing end, the 'middle' and the reducing end of the molecule, respectively. The RMSD data for all 12 simulations are shown in Figures S5–S7 (supplementary material). There are important similarities between the binding mechanisms in trajectories A1 and B1 (Figure 2(a,b)). Most notably, the D and the F rings must be positioned 'correctly' before ring H can reach its binding site found in the X-ray diffraction structures. These results agree with the previous proposal that the DEF part of pentasaccharide molecule binds first, while strong interaction with the GH part requires conformation changes induced in the previous step of the process (Desai et al., 1998). The sequence of events is different in the advanced sampling simulation of the 1NQ9 system (Figure 2(c)). Here, the binding of ring F does not precede ring H, but these steps occurred in the reverse order. The 1NQ9 structure represents an 'intermediate activated' form of AT in which most conformation changes around the heparin-binding site have

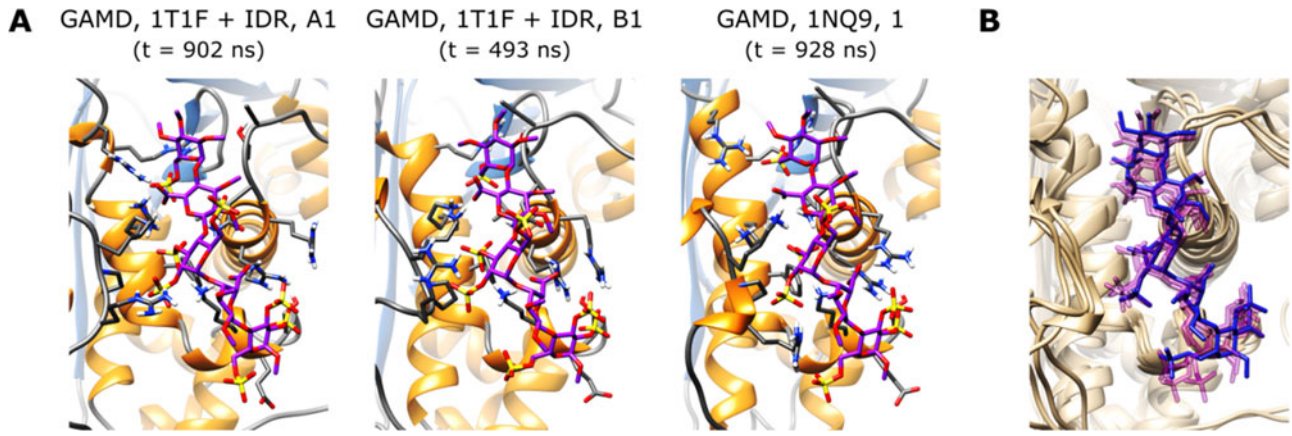

Figure 1. (A) Snapshots taken from the three 1-μs GAMD simulations representing 'low RMSD' conformations of the antithrombin-pentasaccharide complex. (B) The three-binding modes were superimposed. The position of the 'NTP' ligand in the 1NQ9 X-ray diffraction structure is also shown for comparison (*dark blue*).

taken place with the exception of D helix extension. Thus, the binding site for ring H is already in a conformation that can bind the GH end of the ligand with a high affinity.

The mechanism of pentasaccharide binding was further examined by GAMD reweighting analysis. For all GAMD simulations with pentasaccharide, this calculation was performed using three pairs of reaction coordinates: the RMSD values for ring D vs. ring F, ring D vs. ring H and ring F vs. ring H (Figures S8–S10, supplementary material). Although the sampling is probably insufficient for exact free energy calculations, some 'qualitative' conclusions can be still drawn from the data. Most free energy surfaces obtained by reweighting show that large part of the conformational space can be explored by the system without any energy barrier higher than 2 kcal/mol. A possible explanation for the relatively small energy differences is that due to the flexibility of the Arg and Lys residues, electrostatic interactions may still be possible despite the RMSDs of one or more rings are somewhat higher.

Positively charged residues near the heparin binding site of AT play a critical role in the interaction with the pentasaccharide (Olson, Björk, & Bock, 2002; Olson et al., 2010). The three most important residues are Lys114, Lys125 and Arg129, but Lys11, Arg13, Arg46 and Arg47 also make significant contributions (Olson et al., 2002). We analyzed the distances between the negatively charged groups in the pentasaccharide and the positively charged residues of the heparin binding site. A cut-off value of 5 Å was used to discriminate between conformations where the interaction was present or absent (Figures 3 and S11, supplementary material). For the binding data from all 12 simulations, see Figure S11 (supplementary material). Our data demonstrate that the ligand can interact with most basic residues even in conformations where the RMSD is relatively high. Probably, this can be explained by the high flexibility of the Arg and Lys residues. Even if most electrostatic interactions may be presented in the higher RMSD conformations, the interaction energy is probably still lower than in the fully bound state, because the AT-pentasaccharide interaction is known to have a large hydrophobic component, and this component depends on specific orientation of AT residues (Jairajpuri et al., 2003).

The subunit G (iduronic acid) of the heparin pentasaccharide can adapt different conformations: chair and boat-like. To investigate the ring conformations during the simulation, we have computed the Cremer-Pople theta parameter for all five rings (Cremer & Pople, 1975). The 'puckering' data are shown in Figure S12 (supplementary material). Significant conformation changes were observed exclusively in the B1 simulation and only in ring G. However, we have only observed minor conformational flexibility in ring E where an alternative, boat-like form was also present in our pentasaccharide-only GAMD simulations (Balogh et al., 2019). The conformations of rings D, F and H were remarkably stable in all simulations and exclusively chair conformation was observed in all cases.

3.2. Conformational changes of antithrombin

The conformational changes of AT accompanying the pentasaccharide binding have been grouped into two stages. The first set of these changes corresponds to the induced-fit events following the initial weak binding of the ligand. Conformational changes mostly occur near the heparin binding site, the D-helix and the N-terminal part of the A helices. The formation of the P helix near the N-terminal of the D-helix is also considered to take place in the first step (Olson et al., 2010). However, in the 1T1F X-ray diffraction structure of AT, the molecule was captured in a native-like conformation that contained an already formed P helix (Johnson et al., 2006). In the next stage, the extension of the D helix leads to locking of the hinge region in its expelled state. This causes further increases of the heparin binding affinity and protease inhibitory activity of AT.

Modelling of helix P formation is challenging, because the insufficiency of the conformational sampling, even with advanced sampling methods. This is supported by our two 'preliminary' GAMD simulations, based on the 1E04 (McCoy et al., 2003) X-ray structure that does not contain this helix. The DSSP analysis of these simulations showed no sign of helix formation in this position (see Figure S13, supplementary material). As an alternative solution, we have chosen a starting structure of native AT with a pre-formed helix P, 1T1F. To assess the stability of this secondary structural element in our MD and GAMD simulations, we have

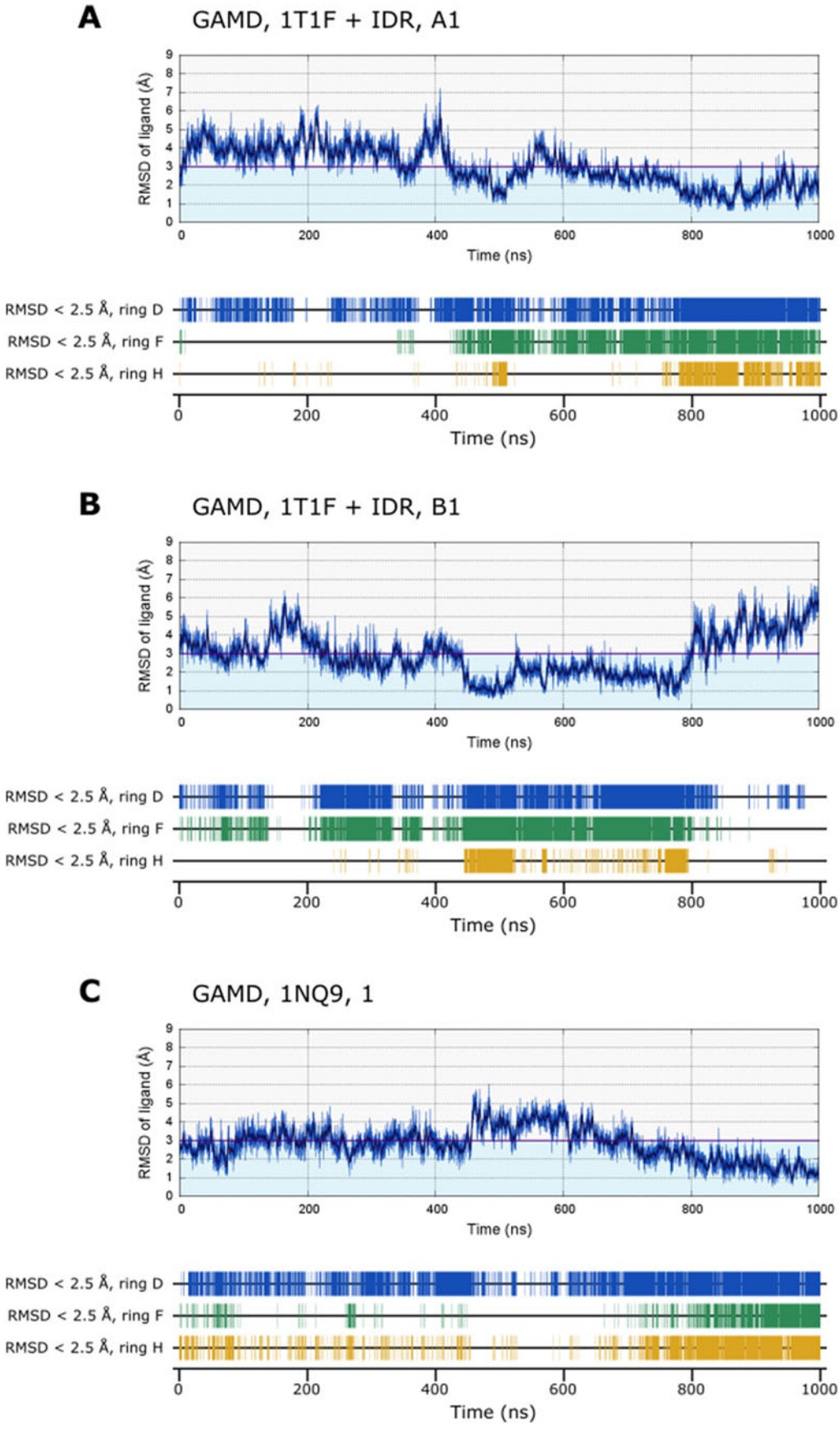

Figure 2. 'Sequential' binding of the D, F, and H rings of the pentasaccharide in the A1 and B1 GAMD simulations. The H ring reaches its final 'position' after the D and F ring has been correctly positioned. The mechanism is different in case of the 1NQ9 simulation. The method of the RMSD calculation for the entire ligand and the specific rings is discussed in detail in the 'Trajectory analysis' section.

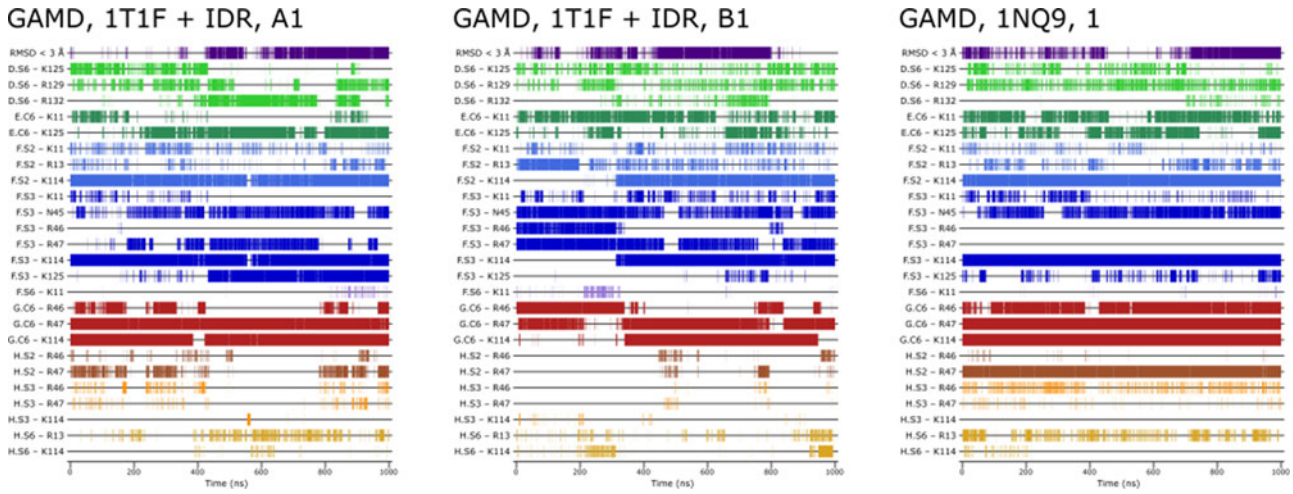

Figure 3. Interactions (distance < 5 Å) between the negatively charged groups of the pentasaccharide and the basic residues of the AT heparin binding site, during the three 1-μs simulations.

performed DSSP analysis of this region. The helix was very stable even in advanced sampling simulations, practically no change occurred in this region in all but one trajectories. The only exception is one of the two GAMD simulations of AT without an activator (Figures 4(a) and S14, supplementary material). Here an alternative conformation of the protein backbone can be observed around Asn117. The DSSP data also indicate that the secondary structure is no longer α-helical in this position.

The possibility of a conformation change near the C terminal of the helix may provide information on the mechanism concerning the formation of this structural element. However, in the N-terminal part (amino acids 112–116) no significant changes can be observed. Therefore, our observations do not contradict our hypothesis of the high stability of this region. We could confirm the high stability of this helix using an advanced sampling technique, proposed previously in a work from our group on the basis of conventional MD simulations (Tóth et al., 2015). The high stability of helix P even in AT conformations without allosteric activation suggests that this form may be in equilibrium with 'native' forms without such a structural element, and a significant fraction of circulating native AT could adapt this conformation. Based on these findings we suggest that ligand binding studies starting from a conformation with this helix may provide biologically relevant data.

Recently, it has been proposed that the hinge region expulsion takes place only in the last step of activation (Olson et al., 2010; Roth et al., 2015). In contrast, Langdown et al. (2009) suggested a rapid equilibrium between the 'inserted' and 'expelled' forms and emphasized the importance of conformational changes of beta-sheet A that result in 'trapping' of hinge region in the 'expelled' state.

The distance between the alpha carbon atoms of Val375 and Ser380 can be used to distinguish between the two conformational types for the hinge region. This parameter is plotted as a function of time in Figure S15 for all 12 simulations. Hinge region expulsion took place in all but two (B2 and one of the 1T1F simulations without ligand) GAMD simulations (and also in one of the cMD simulations of the 1T1F system with a pentasaccharide.). Re-insertion of RCL's N-terminal end for a longer period of time was observed in three trajectories, including the A1 and the expanded 1NQ9 simulation (Figures 5 and S15, supplementary material).

An important question is how the pentasaccharide binding affects the conformation of the hinge region. To determine the relative energies of the various hinge-region inserted, expelled and intermediate states and to study the effects of the ligand binding, we performed GAMD reweighting. Free energies were calculated as a function of two reaction coordinates: the RMSD of the pentasaccharide and the Val375–Ser380 distance describing hinge region conformations (Figures 6(a) and S16, supplementary material). These calculations were performed on all GAMD simulations containing a pentasaccharide. Two further GAMD reweighting calculations were performed that included all data points from the four '1T1F with added pentasaccharide' and the two '1NQ9' GAMD simulations (Figure 6(b)). Our results are compatible with a model in which the hinge region expulsion can take place relatively easily in different states of activation. The closure of the beta-sheet A may be the critical conformation change that is specific to the highest activation state. Such closure of the sheet is evident in case of the 1-μs simulation of the 1NQ9 system in Figure 5(c). As an alternative explanation, AT may have reached a fully activated state with the hinge region expulsion. This seems unlikely in case of the A1 and B1 systems, due to the lack of helix D extension. However, the final 150 ns part of the 1-μs 1NQ9 simulation, where helix extension has taken place, may represent a fully activated state of AT (Figure 4(b)).

Significant C-terminal extension of the D helix was observed in only one trajectory, the 1-μs 1NQ9 simulation (Figure S14, supplementary material). Here, according to the DSSP data (Figure 4(b)) the C-terminal part of D helix did not undergo any significant conformation changes in the first part of the simulation. Significant helix extension could be first observed at around 720 ns. Around the same time, a decrease in the RMSD of the pentasaccharide can be seen in

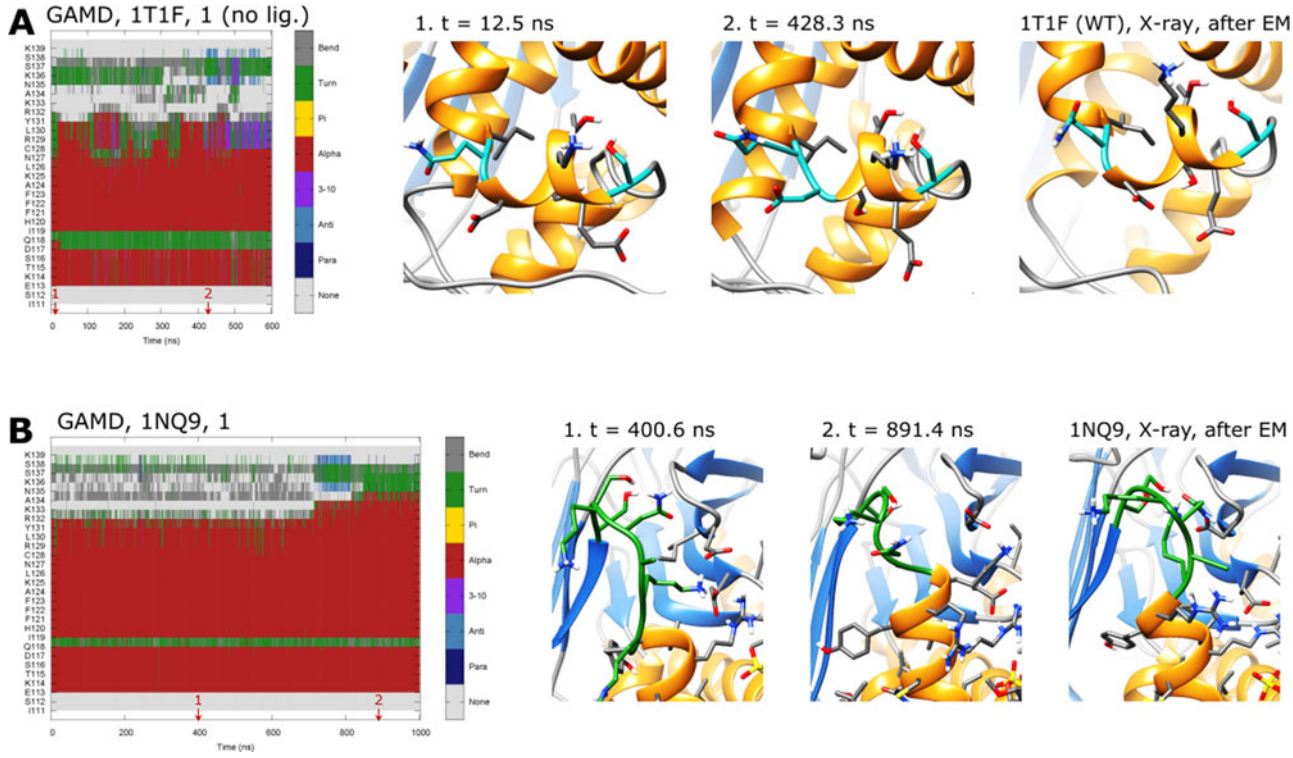

Figure 4. (A) Conformation change of the P helix in one of the GAMD simulations without a pentasaccharide ligand. The DSSP analysis of the secondary structure in the region 110–139 is depicted, along with 'characteristic' conformations from the trajectory (for the method of obtaining the representative structures, see the 'Trajectory analysis' section). The 1T1F X-ray diffraction structure (after modelling the missing loops and energy minimization) is also shown as a comparison. (B) Extension of the D helix in one of the GAMD simulations, starting from the 1NQ9 X-ray diffraction structure. 'Representative' conformations were obtained by means of clustering (see Figure S2A).

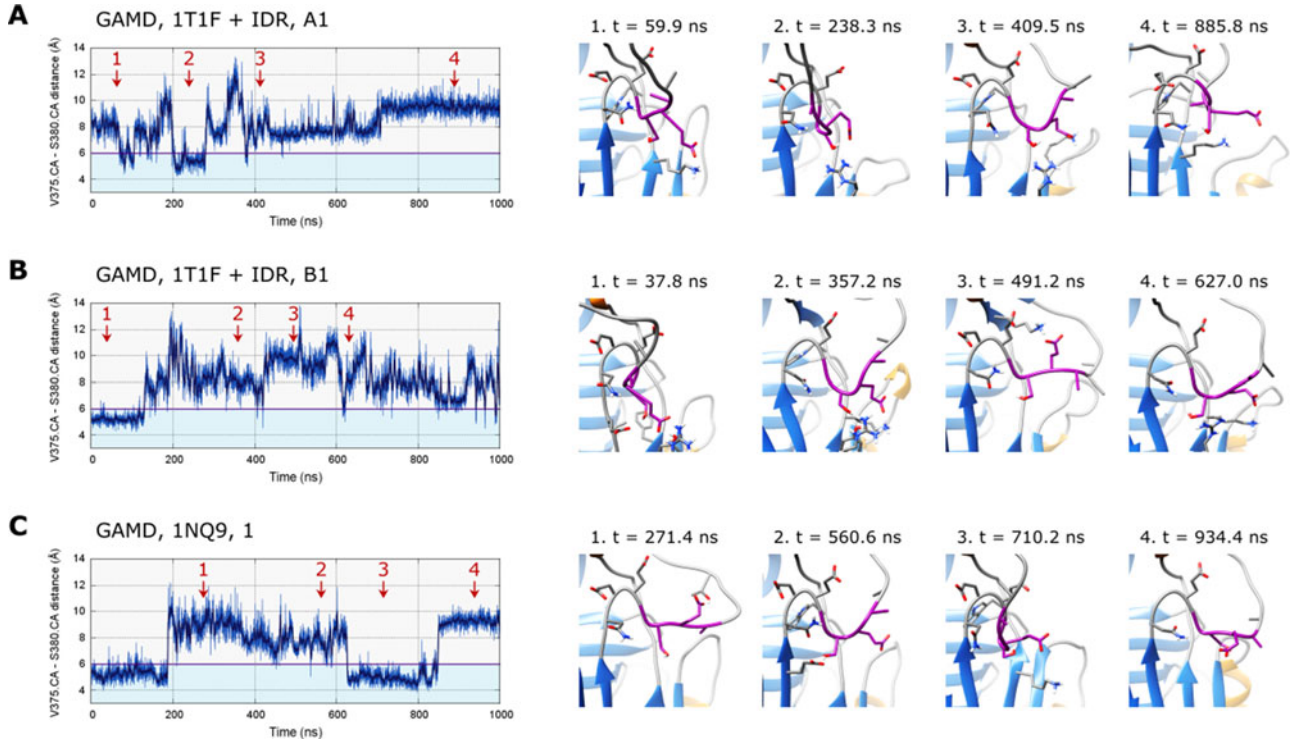

Figure 5. 'Characteristic' conformations of the 'hinge region' (the N-terminal end of the RCL) from the three 1-μs GAMD simulations. The V375.CA–S380.CA distance was used to distinguish between the various conformations. Low values (less than approximately 6 Å) correspond to the 'inserted' state. 'Representative' conformations were obtained by means of clustering (see Figure S2B).

Figure 2c. In the last 150 ns, the Ala134 amino acid was predicted to be part of the helix in around half of the simulation time. This corresponds to a conformation in which the helix is approximately one amino acid shorter than in the X-ray diffraction structure 1E03 (McCoy et al., 2003), representing a 'fully activated' form.

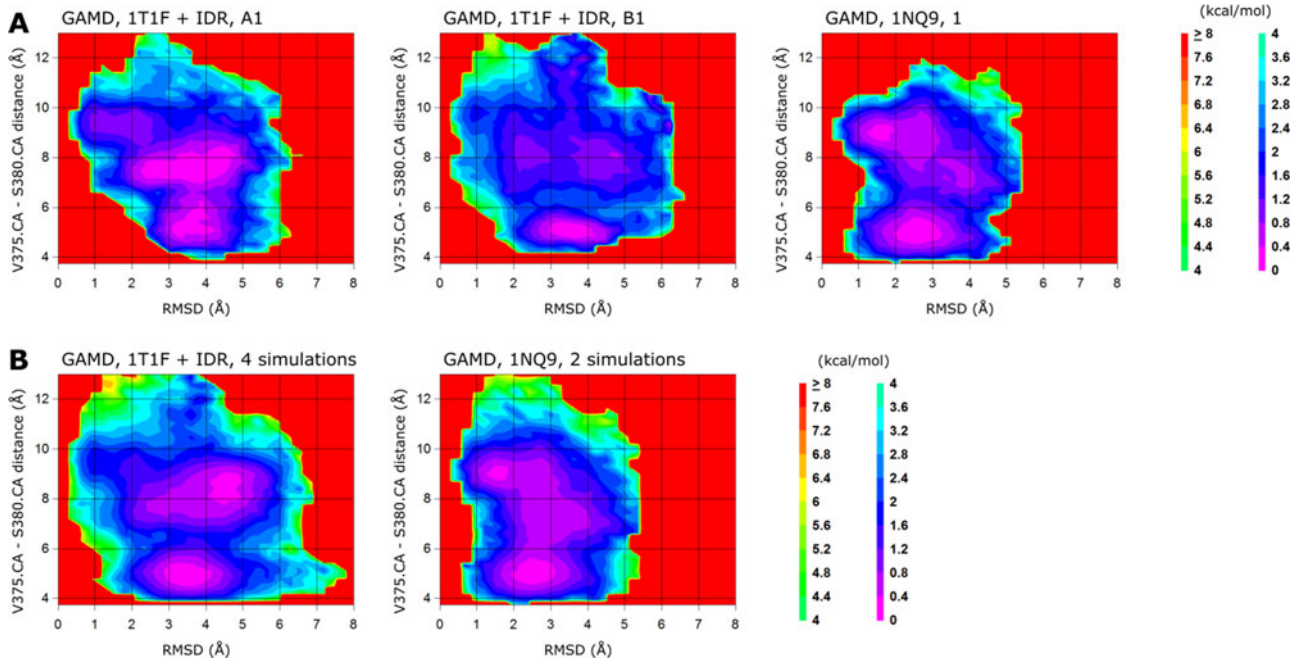

Figure 6. Two-dimensional plots of the results obtained by GAMD reweighting, using two reaction coordinates: RMSD of the pentasaccharide ring and backbone atoms vs. V375.CA–S380.CA distance. (A) The results obtained from the three 1 μs simulations. (B) Two further GAMD reweighting calculations, including all data from the four 1T1F and the two 1NQ9 simulations.

3.3. Conformation of the reactive center loop

The reactive center loop of the inhibitory serpins plays a crucial role in the inactivation of target proteases. The importance of conformational dynamics of this loop in the AT protein has been emphasized in a previous paper form our group (Tóth et al., 2015). A significant number of X-ray diffraction structures are available for the AT protein. However, in many cases, the complex between a native and a latent AT molecule was crystallized and its structure resolved. In these dimeric X-ray diffraction structures, the RCL of the native serpin molecule interacts with the latent AT. Thus, such structures could provide only limited information on the structure of this loop in a solution phase. In a latter X-ray diffraction study, the structure of the native AT was successfully determined that did not interact with a latent serpin molecule. In this experiment, an AT variant with an engineered disulfide bond was used that prevented the latent transition of AT. In this structure (1T1F), the RCL of AT has adopted a novel conformation compared to the previous studies. In a previous work, long (2 μs) molecular dynamics simulations of non-activated AT were performed and analyzed, representing both conformation types (Tóth et al., 2015). In the 'conventional' MD simulation, no transition between the two forms could be observed. However, it was demonstrated using a metadynamics simulation that this is possible in at least one direction (Tóth et al., 2015).

In the present work the Arg236–Ile390 distance was chosen for comparing the various conformations as it can separate the two major forms. In Figure 7, this distance is plotted as a function of time, along with several 'characteristic' conformations from the three 1-μs trajectories. 'Representative' conformations were obtained using clustering (see Figure S2C, supplementary material). The time evolution of the same parameter is presented for all 12 simulations in Figure S17 (supplementary material).

At the beginning of both the A1 and B1 simulations, the distance of the Arg236 and Ile390 alpha carbon atoms is well below ~15 Å in most analyzed conformations. This corresponds to a loop position comparable to that found in the 1T1F structure. At around 280 ns in the A1 and around 140 ns in the B2 simulation, a remarkable conformation change occurs. From that point, in the A1 simulation various RCL conformations can be observed, and the distance parameter fluctuates remarkably. The distance values are approximately in the same range as in the 1NQ9 trajectory, derived from an X-ray structure with a different starting RCL conformation. In contrast, the same parameter was relatively stable in the last 800 ns part of the simulation B1. Regarding the 1 μs simulation based on the 1NQ9 structure, the existence of large number of different RCL conformations can be seen. However, conformations in the 15–18 Å region were only minimally sampled, and no RCL position resembling to that found in the 1T1F structure could be observed. A solvent accessible Arg393 residue, required for efficient inhibition of target proteases, was present in a significant number of structures from all three simulations.

However, conformational changes between the two major conformational types are still 'rare events' even in the advanced sampling simulations. In all trajectories analyzed, either a single transition occurred in only one direction or no such change was observed at all. To obtain a rough estimation of the potential surface of the RCL conformations GAMD reweighting calculations were performed (Figure S18, supplementary material), using two reaction coordinates, the conformation of RCL and of the hinge region. Although – due to insufficient convergence – there is uncertainty in the reweighting data, the results are compatible with an energy

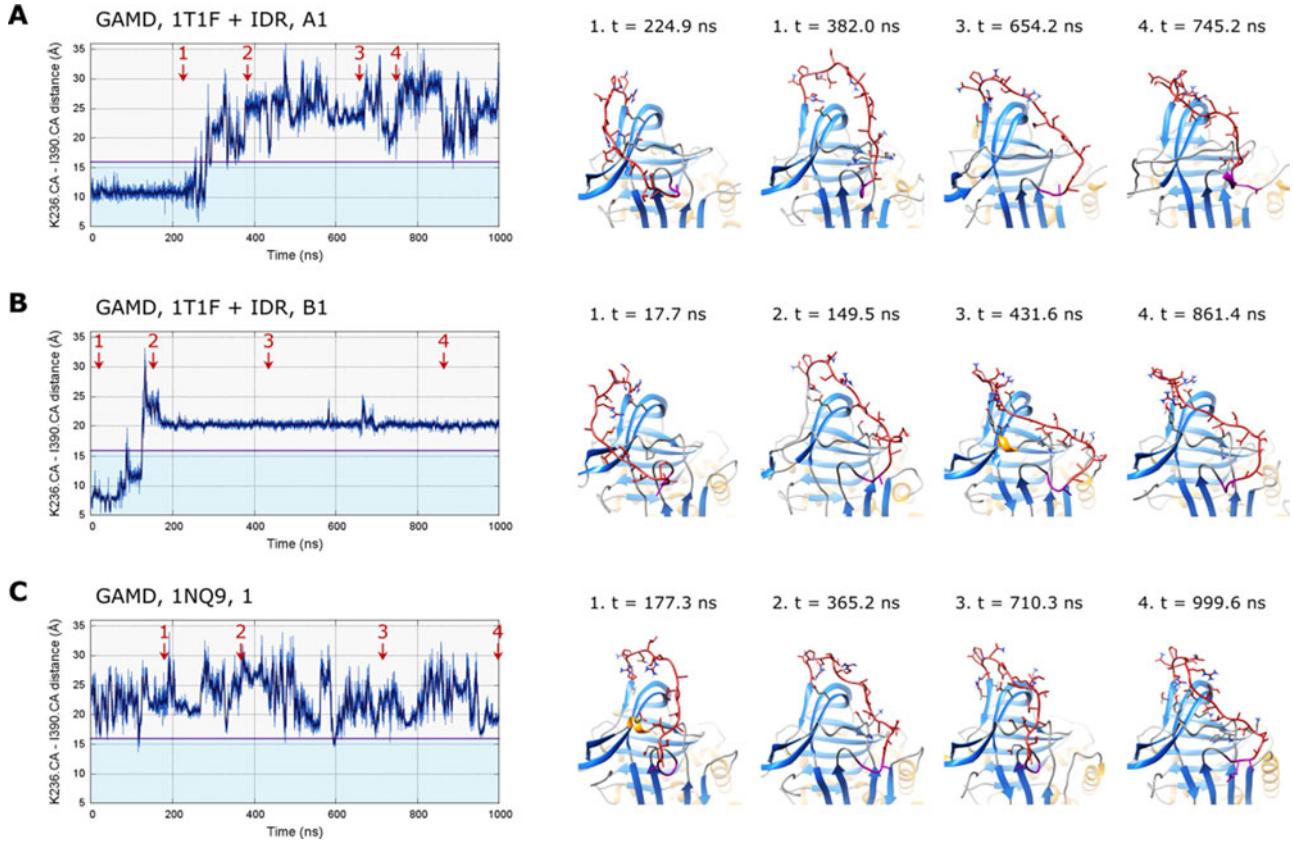

Figure 7. 'Characteristic' conformations of the reactive center loop from the three 1-μs GAMD simulations. The K236.CA–I390.CA distance was used to distinguish between the various conformations. Low values (5–16 Å) represent conformations similar to that found in the 1T1F X-ray structure. 'Representative' conformations were obtained by means of clustering (see Figure S2C).

barrier in the 15–20 Å region. This energy barrier might slow down the 'mixing' of the two conformational types (Figure S18, supplementary material). Despite of the advanced sampling method by significant computational efforts, however, the conformational sampling could have been still insufficient.

3.4. Root mean square fluctuations

Root mean square fluctuation (RMSF) is a frequently used method to measure flexibility of various structural regions in proteins from an MD trajectory. To study the changes in fluctuation caused by the pentasaccharide binding, we have performed RMSF calculations on all 12 'production' simulations (Figures 8 and S19–S21, supplementary material).

As reported previously (Tóth et al., 2015), most AT residues have relatively low root mean square fluctuations, especially those that form helices or beta-sheets. Notable exceptions are the RCL and the loops in the N-terminal region. These loops are highly flexible, in agreement with the previously published RMSF data from conventional MD simulations. Other regions with higher than average RMSF values are the loop involved in the D helix extension (residues 131–139), the loop at the C terminal end of helix D (residues 191–210), some of the residues close to the FXa binding site (amino acids 235–240) and a flexible segment consisting of amino acids 355–365.

To investigate the effects of pentasaccharide binding on the fluctuations, another set of RMSF calculations was carried out for the three 1-μs simulations. Here only those parts were included in the calculations where most of the conformations with low pentasaccharide RMSD occurred: 640–1000 ns for the A1, 440–800 ns for A2 and 720–1000 ns for the 1-μs simulation of the 1NQ9 system (compare Figure 2(a)–(c)).

Decreased fluctuations can be observed near the heparin binding site (residues 110–140) in both the A1 and A2 simulations. The decrease of RMSF near helices D and P may suggest the stabilization of a conformation, which is more suitable for binding the ligand. Further changes in RMSF can be observed in case of residues 220–240. These changes are more significant in case of the A1 simulation, but are also clearly present in case of simulation B1. This region contains multiple amino acids with a proposed role in AT-FXa exosite interactions, namely Asn233, Arg235 and Glu237 (Izaguirre et al., 2014). The decrease of RCL fluctuations may be surprising as increased RCL flexibility triggered by ligand binding has been reported in the literature (Arantes et al., 2018). The decrease of RCL fluctuation observed in our study could be explained by the fact that RCL was highly flexible in our simulations and several of these conformations were excluded from the analysis. The RCL remains, however one of the most flexible structural elements of the protein.

3.5. Mechanism of allosteric activation

The investigation of allosteric pathways are essential for understanding the conformational activation of AT. Multiple proposals of possible pathways exist in literature. Both X-ray

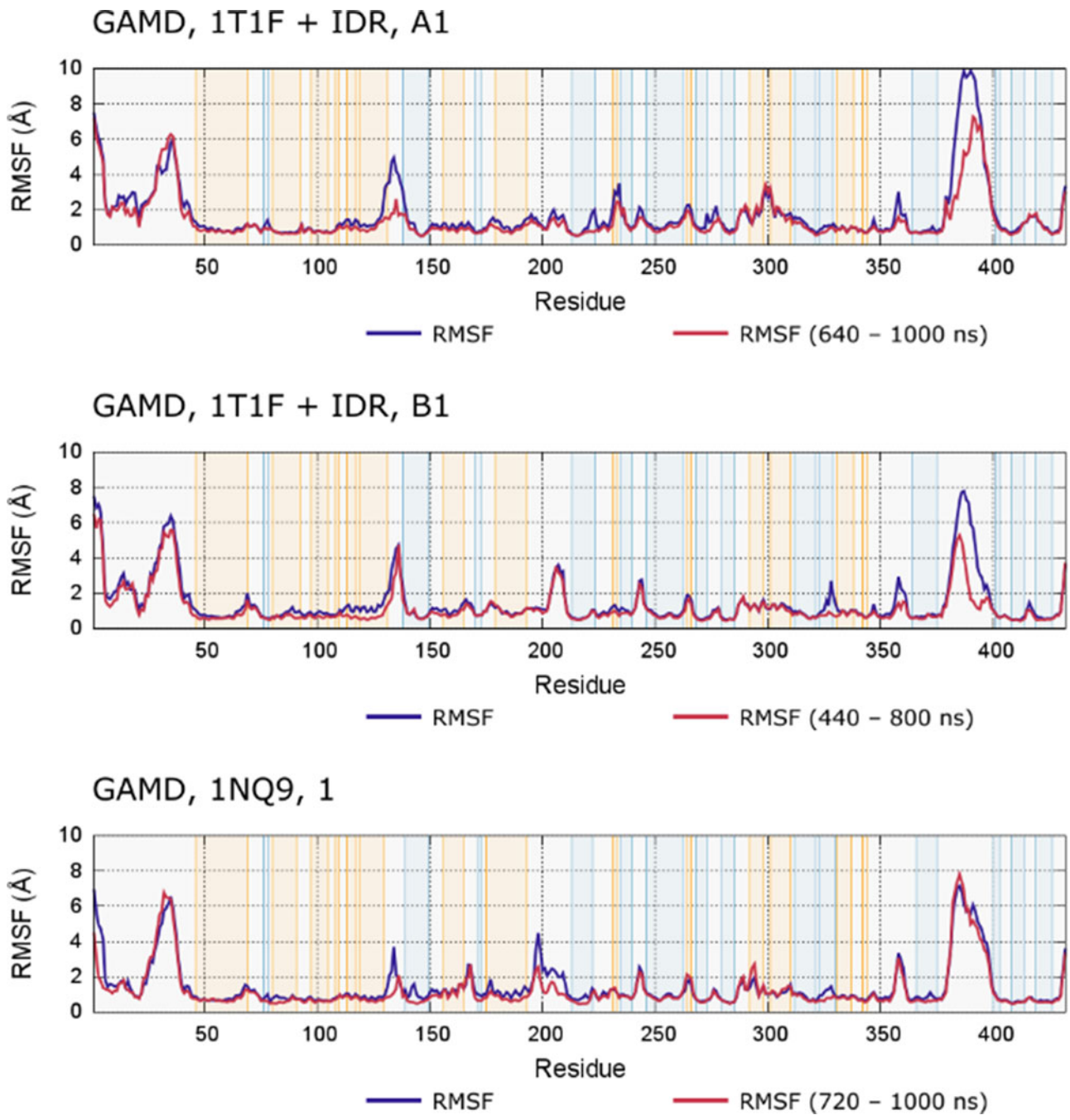

Figure 8. Root mean square fluctuations (RMSF) of the α-carbon atoms, during the three 1 μs molecular dynamics simulations. The RMSF values calculated for the whole trajectory are shown in blue. RMSF was also computed for parts of the trajectories where low-RMSD binding states of the pentasaccharide were most 'typical' (depicted in red). The backgrounds of the plots were colored according to the secondary structure of the region: orange for helices and blue for β-sheets.

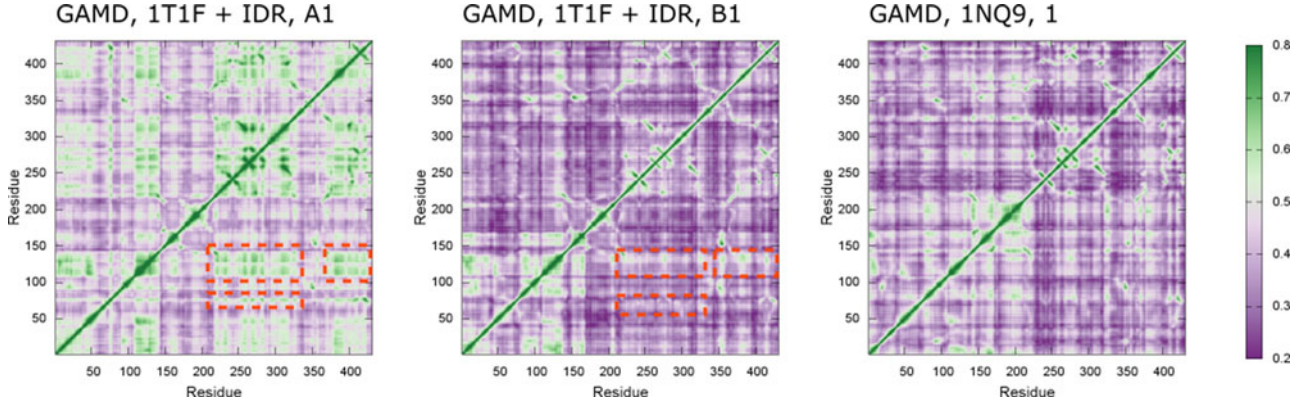

Figure 9. Generalized correlation matrices for the three 1-μs GAMD trajectories, calculated using a method proposed by Lange and Grubmüller (2005).

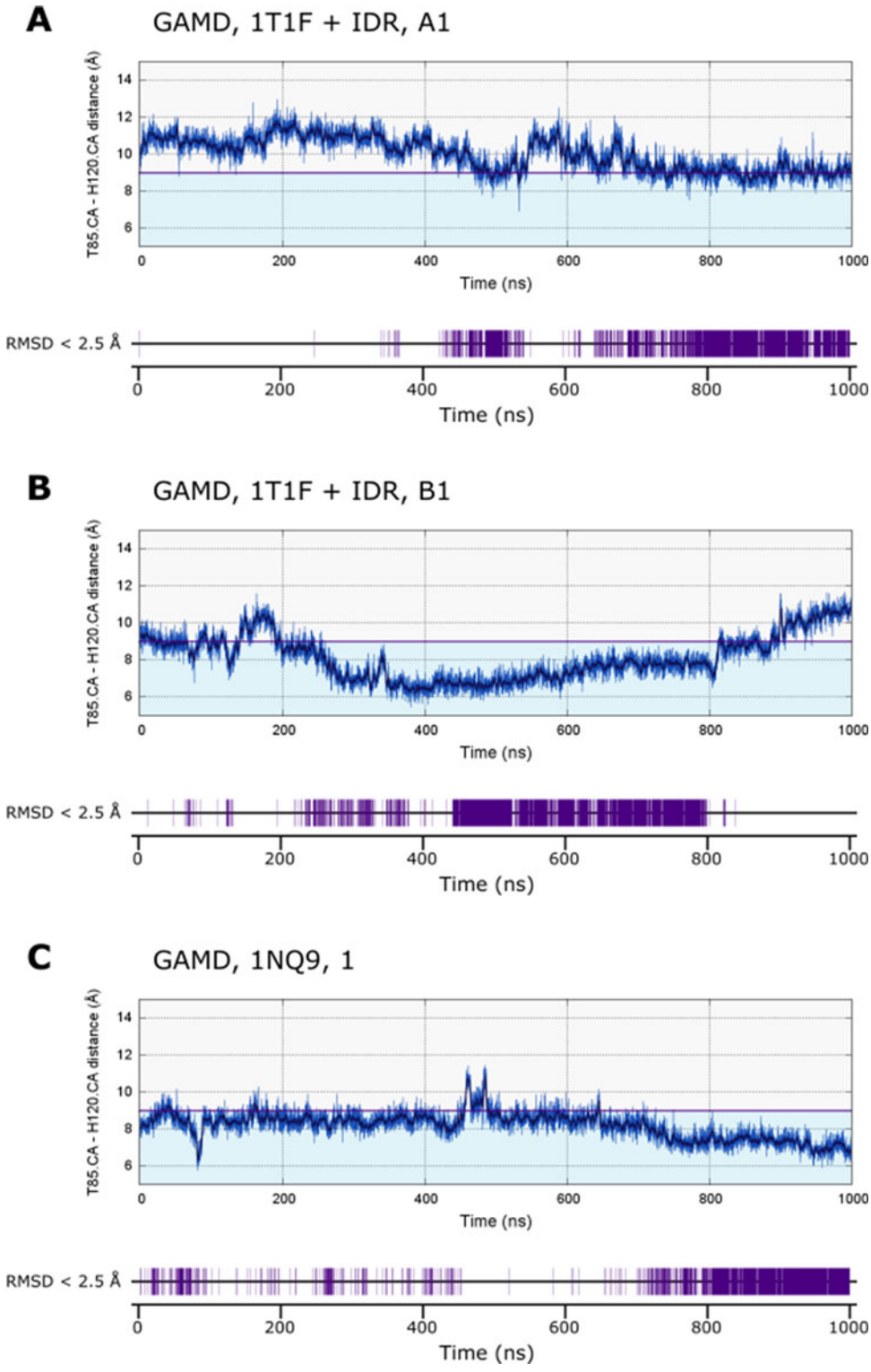

Figure 10. A conformation change involving the N-terminal part of the D helix. The distance between the alpha carbon atoms of Thr85 (in helix B) and His120 correlates with the binding mode of the pentasaccharide.

diffraction structures (Huntington, 2011; Olson et al., 2010) and MD-based studies (Arantes et al., 2018; Pol-Fachin et al., 2011; Tóth et al., 2015) have been used to investigate the details of the allosteric processes. In our GAMD simulations, we were able to capture a significant number of different binding modes for the pentasaccharide, including conformations with relatively large RMSD compared to the X-ray structures. Therefore, it can be assumed that such advanced sampling simulations may potentially reveal details of the allosteric processes that would be difficult to obtain using either X-ray crystallography data or 'conventional' MD simulations.

From the molecular dynamics trajectories, we analyzed the correlated motions between the alpha carbon atoms using the 'generalized correlation' or 'mutual information' measure, developed by Lange and Grubmüller (2006). A

significant advantage of this method over the probably more common Pearson correlation coefficient calculations is that it can also detect nonlinear correlations between specific atoms of the proteins. The 'generalized correlation' matrices, calculated for all 12 'production' simulations, are depicted in Figure S22 (supplementary material). The matrices for the A1 and the B1 systems as well as the 1 μs 1NQ9 simulations are compared in Figure 9. Regarding the A1 simulation, interesting correlated motions can be seen between two distant parts of the protein: helix D and adjacent residues, critical for heparin binding and the region consisting of the 230–310 amino acids. Many known exosite residues involved in FXa and FIXa binding are found in the latter region: Asn233, Arg235, Glu237, Tyr253 and Glu255 (Izaguirre et al., 2014; Olson et al., 2010). Correlation was also detected between helix D and the N-terminal part of reactive center loop, confirming the relationship between binding of the pentasaccharide, hinge region expulsion and conformational changes of RCL. A further region of AT that seems to be involved in the allosteric processes consists of amino acids 75–78. These residues are located between the helices A and B and belong to the hydrophobic core of the molecule. The C-terminal part of the molecule (residues ~400–432) shows similar correlations with the factor-binding region, which can be possibly explained by the proximity of these residues to both the C-terminal end of helix D and the FXa binding exosite. Most other correlated motions are found between regions that are close to each other in the tertiary structure. Regarding the B1 simulation, correlations can be seen in approximately the same regions, although the coefficients are lower and a smaller part of the D helix is involved. Rather than amino acids 75–78, the C-terminal part of helix A (residues 67–70) shows correlated motions.

In contrast, the 'mutual information' matrix computed for the 1-μs 1NQ9 simulations show significant differences compared to the previous two, although there are also similarities. Here, correlated motions can be observed between the previously mentioned exosite, the RCL, the C-terminal part of the protein as well as most amino acids in the region 110–230. This latter region contains helices D, E and F as well as two strands of beta-sheet A. Residues 75–98, containing helix B also seems to be involved in the allosteric processes in this activation state of AT.

A notable conformational change that shows significant correlation with the ligand RMSD, involves the N-terminal part of helix D and helix B. The distances between the alpha carbon atoms of Thr85 (helix B) and His120 (helix D) are shown in Figures 10 and S23 (supplementary material). Helix B occupies a central position in the tertiary structure of AT. It is surrounded by parts of the A, C and D helices and multiple strands in the 'A' beta sheet. Therefore, it could play a significant role in the allosteric processes (Figure 10).

In summary, our 'generalized correlation' calculations revealed possible paths of allosteric signal propagation to the RCL and the binding sites of target proteases. Using the pentasaccharide-binding trajectories, we could detect correlations involving the helix D that were not present in previously published 'conventional' MD simulations. Based on our calculations, we can suggest that parts of the A helix and the C-terminal end of the protein may be involved in the allosteric processes. Helix B may also contribute to the allosteric activation of AT.

4. Conclusion

In the present study, we have studied the early events of pentasaccharide ligand binding to AT using MD and GAMD simulations. The binding of a model compound, idraparinux, was investigated to an X-ray diffraction structure of a 'native' AT conformation without a heparin-like activator. We were able to observe binding modes for the model ligand similar to that found in the 'intermediate' activated AT structure 1NQ9. Such binding occurred in two independent GAMD trajectories, with RMSD values as low as 0.6–0.8 Å for the ring and interglycosidic atoms of the molecule. Furthermore, we analyzed various binding positions of pentasaccharide in multiple MD trajectories, including some, where the 'intermediate' activated AT structure was selected as starting conformation. Our simulations also provided insights into conformational changes of the hinge region, the RCL, and the C-terminal part of helix D. The observed high stability of helix P might indicate that AT conformations with an already formed P helix could play a significant role in the pentasaccharide binding and conformational activation. Our 'generalized correlation' calculations provided us details on possible allosteric pathways. However, the modelling of the complex process of the induced-fit heparin pentasaccharide binding remains a challenging task, even with advanced sampling techniques such as GAMD.

In summary, our results demonstrate the applicability of the GAMD method for investigating the binding of a heparin pentasaccharide to AT. This ligand has several characteristics that distinguish it from the molecules applied earlier in related binding studies: it is relatively large, highly flexible due to its carbohydrate-like structure, and highly negatively charged. Regarding the potential applications of this method, future binding studies with various pentasaccharide or even non-pentasaccharide AT activators could aid the design of novel AT-dependent anticoagulant drugs. Furthermore, the biochemical consequences of several pathogenic variants affecting the heparin binding of AT can be investigated by this method (Corral, de la Morena-Barrio, & Vicente, 2018; Dinarvand et al., 2018; Muszbek et al., 2010). Finally, such simulations could potentially serve as an in silico tool for investigating interactions between numerous GAG-binding proteins with diverse biological functions and sulfated glycosaminoglycans.

Acknowledgements

We are thankful to Prof. László Muszbek for his useful comments and suggestions. We acknowledge Governmental Information Technology Development Agency for awarding us access to supercomputing resources based in Debrecen, Hungary. Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Hungarian Scientific Research Fund (OTKA K-106294, 116228 and 120633) and by the Ministry of National Economy, Hungary (GINOP-2.3.2-15-2016-00039 and GINOP-2.3.2-15-2016-00050). Gabor Balogh is a recipient of a scholarship from the "Talentum" Foundation of Gedeon Richter PLC.

References

Andersen, O. J., Risør, M. W., Poulsen, E. C., Nielsen, N. C., Miao, Y., Enghild, J. J., & Schiøtt, B. (2017). Reactive center loop insertion in α-1-antitrypsin captured by accelerated molecular dynamics simulation. *Biochemistry*, *56*(4), 634–646. doi:10.1021/acs.biochem.6b00839

Arantes, P. R., Pérez-Sánchez, H., & Verli, H. (2018). Antithrombin conformational modulation by D-myo-inositol 3,4,5,6-tetrakisphosphate (TMI), a novel scaffold for the development of antithrombotic agents. *Journal of Biomolecular Structure and Dynamics*, *36*(15), 4045–4056. doi:10.1080/07391102.2017.1407259

Babin, V., Karpusenka, V., Moradi, M., Roland, C., & Sagui, C. (2009). Adaptively biased molecular dynamics: An umbrella sampling method with a time-dependent potential. *International Journal of Quantum Chemistry*, *109*(15), 3666–3678. doi:10.1002/qua.22413

Balogh, G., Gyöngyösi, T., Timári, I., Herczeg, M., Borbás, A., Fehér, K., & Kövér, K. E. (2019). Comparison of carbohydrate force fields using Gaussian Accelerated Molecular Dynamics simulations and development of force field parameters for heparin-analogue pentasaccharides. *Journal of Chemical Information and Modeling*. doi:10.1021/acs.jcim.9b00666

Barducci, A., Bonomi, M., & Parrinello, M. (2011). Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(5), 826–843. doi:10.1002/wcms.31

Belzar, K. J., Zhou, A., Carrell, R. W., Gettins, P. G., & Huntington, J. A. (2002). Helix D elongation and allosteric activation of antithrombin. *Journal of Biological Chemistry*, *277*(10), 8551–8558. doi:10.1074/jbc.M110807200

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690. doi:10.1063/1.448118

Bernardi, R. C., Melo, M. C. R., & Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (Bba) - General Subjects*, *1850*(5), 872–877. doi:10.1016/j.bbagen.2014.10.019

Bhattarai, A., & Miao, Y. (2018). Gaussian accelerated molecular dynamics for elucidation of drug pathways. *Expert Opinion on Drug Discovery*, *13*(11), 1055–1065. doi:10.1080/17460441.2018.1538207

Brennan, S. O., George, P. M., & Jordan, R. E. (1987). Physiological variant of antithrombin-III lacks carbohydrate sidechain at Asn 135. *FEBS Letters*, *219*(2), 431–436. doi:10.1016/0014-5793(87)80266-1

Brünger, A. T., Adams, P. D., & Rice, L. M. (1997). New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure (London, England: 1993)*, *5*(3), 325–336. doi:10.1016/s0969-2126(97)00190-1

Case, D., Betz, R., Cerutti, D. S., Cheatham, T., Darden, T., Duke, R., … Kollman, P. A. (2016). *Amber 2016*. San Francisco, CA: University of California.

Cazzolli, G., Wang, F., A Beccara, S., Gershenson, A., Faccioli, P., & Wintrode, P. L. (2014). Serpin latency transition at atomic resolution. *Proceedings of the National Academy of Sciences*, *111*(43), 15414–15419. doi:10.1073/pnas.1407528111

Corral, J., de la Morena-Barrio, M. E., & Vicente, V. (2018). The genetics of antithrombin. *Thrombosis Research*, *169*, 23–29. doi:10.1016/j.thromres.2018.07.008

Cremer, D., & Pople, J. A. (1975). General definition of ring puckering coordinates. *Journal of the American Chemical Society*, *97*(6), 1354–1358. doi:10.1021/ja00839a011

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089–10092. doi:10.1063/1.464397

Desai, U. R., Petitou, M., Björk, I., & Olson, S. T. (1998). Mechanism of heparin activation of antithrombin: Role of individual residues of the pentasaccharide activating sequence in the recognition of native and activated states of antithrombin. *Journal of Biological Chemistry*, *273*(13), 7478–7487. doi:10.1074/jbc.273.13.7478

Dinarvand, P., Yang, L., Villoutreix, B. O., & Rezaie, A. R. (2018). Expression and functional characterization of two natural heparin-binding site variants of antithrombin. *Journal of Thrombosis and Haemostasis*, *16*(2), 330–341. doi:10.1111/jth.13920

Guvench, O., Mallajosyula, S. S., Raman, E. P., Hatcher, E., Vanommeslaeghe, K., Foster, T. J., … MacKerell, A. D. (2011). CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *Journal of Chemical Theory and Computation*, *7*(10), 3162–3180. doi:10.1021/ct200328p

Hamelberg, D., Mongan, J., & McCammon, J. A. (2004). Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *Journal of Chemical Physics.*, *120*(24), 11919–11929. doi:10.1063/1.1755656

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., … MacKerell, A. D. (2017). CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nature Methods*, *14*(1), 71–73. doi:10.1038/nmeth.4067

Huntington, J. A. (2006). Shape-shifting serpins–advantages of a mobile mechanism. *Trends in Biochemical Sciences*, *31*(8), 427–435. doi:10.1016/j.tibs.2006.06.005

Huntington, J. A. (2011). Serpin structure, function and dysfunction. *Journal of Thrombosis and Haemostasis*, *9*(Suppl. 1), 26–34. doi:10.1111/j.1538-7836.2011.04360.x

Ichiye, T., & Karplus, M. (1991). Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Structure, Function, and Genetics*, *11*(3), 205–217. doi:10.1002/prot.340110305

Izaguirre, G., Aguila, S., Qi, L., Swanson, R., Roth, R., Rezaie, A. R., … Olson, S. T. (2014). Conformational activation of antithrombin by heparin involves an altered exosite interaction with protease. *Journal of Biological Chemistry*, *289*(49), 34049–34064. doi:10.1074/jbc.M114.611707

Jairajpuri, M. A., Lu, A., Desai, U., Olson, S. T., Bjork, I., & Bock, S. C. (2003). Antithrombin III phenylalanines 122 and 121 contribute to its high affinity for heparin and its conformational activation. *Journal of Biological Chemistry*, *278*(18), 15941–15950. doi:10.1074/jbc.M212319200

Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, *29*(11), 1859–1865. doi:10.1002/jcc.20945

Johnson, D. J., & Huntington, J. A. (2003). Crystal structure of antithrombin in a heparin-bound intermediate state. *Biochemistry*, *42*(29), 8712–8719. doi:10.1021/bi034524y

Johnson, D. J., Langdown, J., Li, W., Luis, S. A., Baglin, T. P., & Huntington, J. A. (2006). Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation. *Journal of Biological Chemistry*, *281*(46), 35478–35486. doi:10.1074/jbc.M607204200

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, *22*(12), 2577–2637. doi:10.1002/bip.360221211

Kästner, J. (2011). Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(6), 932–942. doi:10.1002/wcms.66

Langdown, J., Belzar, K. J., Savory, W. J., Baglin, T. P., & Huntington, J. A. (2009). The critical role of hinge-region expulsion in the induced-fit

heparin binding mechanism of antithrombin. *Journal of Molecular Biology*, *386*(5), 1278–1289. doi:10.1016/j.jmb.2009.01.028

Lange, O. F., & Grubmüller, H. (2005). Generalized correlation for biomolecular dynamics. *Proteins: Structure, Function, and Bioinformatics*, *62*(4), 1053–1061. doi:10.1002/prot.20784

Le Grand, S., Götz, A. W., & Walker, R. C. (2013). SPFP: Speed without compromise – A mixed precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications*, *184*(2), 374–380. doi:10.1016/j.cpc.2012.09.022

Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., … Im, W. (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of Chemical Theory and Computation*, *12*(1), 405–413. doi:10.1021/acs.jctc.5b00935

Liao, J. M., & Wang, Y. T. (2019). In silico studies of conformational dynamics of Mu opioid receptor performed using gaussian accelerated molecular dynamics. *Journal of Biomolecular Structure and Dynamics*, *37*(1), 166–177. doi:10.1080/07391102.2017.1422025

Mallajosyula, S. S., Guvench, O., Hatcher, E., & Mackerell, A. D. (2012). CHARMM additive all-atom force field for phosphate and sulfate linked to carbohydrates. *Journal of Chemical Theory and Computation*, *8*(2), 759–776. doi:10.1021/ct200792v

McCoy, A. J., Pei, X. Y., Skinner, R., Abrahams, J. P., & Carrell, R. W. (2003). Structure of beta-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity. *Journal of Molecular Biology*, *326*(3), 823–833. doi:10.1016/s0022-2836(02)01382-7

Miao, Y., Bhattarai, A., Nguyen, A. T. N., Christopoulos, A., & May, L. T. (2018). Structural basis for binding of allosteric drug leads in the adenosine A_1 receptor. *Scientific Reports*, *8*(1), 16836. doi:10.1038/s41598-018-35266-x

Miao, Y., Feher, V. A., & McCammon, J. A. (2015). Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *Journal of Chemical Theory and Computation*, *11*(8), 3584–3595. doi:10.1021/acs.jctc.5b00436

Miao, Y., Huang, Y. M., Walker, R. C., McCammon, J. A., & Chang, C. A. (2018). Ligand binding pathways and conformational transitions of the HIV protease. *Biochemistry*, *57*(9), 1533–1541. doi:10.1021/acs.biochem.7b01248

Miao, Y., & McCammon, J. A. (2016). Graded activation and free energy landscapes of a muscarinic G-protein-coupled receptor. *Proceedings of the National Academy of Sciences*, *113*(43), 12162–12167. doi:10.1073/pnas.1614538113

Miao, Y., Sinko, W., Pierce, L., Bucher, D., Walker, R. C., & McCammon, J. A. (2014). Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *Journal of Chemical Theory and Computation*, *10*(7), 2677–2689. doi:10.1021/ct500090q

Muszbek, L., Bereczky, Z., Kovács, B., & Komáromi, I. (2010). Antithrombin deficiency and its laboratory diagnosis. *Clinical Chemistry and Laboratory Medicine*, *48*(Suppl 1), S67–S78. doi:10.1515/CCLM.2010.368

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open babel: An open chemical toolbox. *Journal of Cheminformatics*, *3*, 33. doi:10.1186/1758-2946-3-33

Olson, S. T., Björk, I., & Bock, S. C. (2002). Identification of critical molecular interactions mediating heparin activation of antithrombin: Implications for the design of improved heparin anticoagulants. *Trends in Cardiovascular Medicine*, *12*(5), 198–205. doi:10.1016/S1050-1738(02)00160-3

Olson, S. T., Richard, B., Izaguirre, G., Schedin-Weiss, S., & Gettins, P. G. (2010). Molecular mechanisms of antithrombin-heparin regulation of blood clotting proteinases. A paradigm for understanding proteinase regulation by serpin family protein proteinase inhibitors. *Biochimie*, *92*(11), 1587–1596. doi:10.1016/j.biochi.2010.05.011

Park, S. J., Lee, J., Patel, D. S., Ma, H., Lee, H. S., Jo, S., & Im, W. (2017). Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank. *Bioinformatics*, *33*(19), 3051–3057. doi:10.1093/bioinformatics/btx358

Petitou, M., & van Boeckel, C. A. (2004). A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next?. *Angewandte Chemie International Edition*, *43*(24), 3118–3133. doi:10.1002/anie.200300640

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera–A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. doi:10.1002/jcc.20084

Pol-Fachin, L., Franco Becker, C., Almeida Guimarães, J., & Verli, H. (2011). Effects of glycosylation on heparin binding and antithrombin activation by heparin. *Proteins: Structure, Function, and Bioinformatics*, *79*(9), 2735–2745. doi:10.1002/prot.23102

Quinsey, N. S., Greedy, A. L., Bottomley, S. P., Whisstock, J. C., & Pike, R. N. (2004). Antithrombin: In control of coagulation. *The International Journal of Biochemistry & Cell Biology*, *36*(3), 386–389. doi:10.1016/s1357-2725(03)00244-9

Rau, J. C., Beaulieu, L. M., Huntington, J. A., & Church, F. C. (2007). Serpins in thrombosis, hemostasis and fibrinolysis. *Journal of Thrombosis and Haemostasis*, *5*(Suppl 1), 102–115. doi:10.1111/j.1538-7836.2007.02516.x

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, *9*(7), 3084–3095. doi:10.1021/ct400341p

Roth, R., Swanson, R., Izaguirre, G., Bock, S. C., Gettins, P. G., & Olson, S. T. (2015). Saturation mutagenesis of the antithrombin reactive center loop P14 residue supports a three-step mechanism of heparin allosteric activation involving intermediate and fully activated states. *Journal of Biological Chemistry*, *290*(47), 28020–28036. doi:10.1074/jbc.M115.678839

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327–341. doi:10.1016/0021-9991(77)90098-5

Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, *234*(3), 779–815. doi:10.1006/jmbi.1993.1626

Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle Mesh Ewald. *Journal of Chemical Theory and Computation*, *9*(9), 3878–3888. doi:10.1021/ct400314y

Schedin-Weiss, S., Richard, B., & Olson, S. T. (2010). Kinetic evidence that allosteric activation of antithrombin by heparin is mediated by two sequential conformational changes. *Archives of Biochemistry and Biophysics*, *504*(2), 169–176. doi:10.1016/j.abb.2010.08.021

Sugita, Y., & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, *314*(1-2), 141–151. doi:10.1016/S0009-2614(99)01123-9

Tóth, L., Fekete, A., Balogh, G., Bereczky, Z., & Komáromi, I. (2015). Dynamic properties of the native free antithrombin from molecular dynamics simulations: Computational evidence for solvent- exposed Arg393 side chain. *Journal of Biomolecular Structure and Dynamics*, *33*(9), 2023–2036. doi:10.1080/07391102.2014.986525

Wang, F., Orioli, S., Ianeselli, A., Spagnolli, G., A Beccara, S., Gershenson, A., … Wintrode, P. L. (2018). All-atom simulations reveal how single-point mutations promote serpin misfolding. *Biophysical Journal.*, *114*(9), 2083–2094. doi:10.1016/j.bpj.2018.03.027

Whisstock, J. C., & Bottomley, S. P. (2006). Molecular gymnastics: Serpin structure, folding and misfolding. *Current Opinion in Structural Biology*, *16*(6), 761–768. doi:10.1016/j.sbi.2006.10.005